Mumbai University

Question Paper

[IDOL – Revised Course]

(May – 2016)

PAPER - III

Data Warehousing

Time: 3 Hours **Total Marks:** 100

N.B.: **(1)** *All Question are Compulsory.*
(2) *Make Suitable Assumptions Wherever Necessary And State The Assumptions Made.*
(3) *Answer To The Same Question Must Be Written Together.*
(4) *Number To The Right Indicates Marks.*
(5) *Draw Neat Labeled Diagrams Wherever Necessary.*
(6) *Use of Non – Programmable Calculator is allowed.*

Q.1 ATTEMPT ANY TWO QUESTIONS: (10 MARKS)

(A) Explain the term Data Warehouse. **(5)**
(B) What is the importance of metadata in the Data Warehouse? **(5)**
(C) What are the two common types of transactions in Data Warehousing? **(5)**
(D) How the data is protected in Data Warehouse? **(5)**

Q.2 ATTEMPT ANY THREE QUESTIONS: (15 MARKS)

(A) Discuss the impact of the Data Warehouse on Business. **(5)**
(B) Discuss the problems related with Federated Data Warehouse. **(5)**
(C) Discuss the Fundamental Operating differences between the various sectors in Data Warehouse. **(5)**
(D) Explain by giving example structured and unstructured data with respect to a Data Warehouse. **(5)**
(E) What is the significance of referential integrity of the data in the Data Warehouse? **(5)**
(F) The challenges of incorporating unstructured data with structured data in a Data Warehouse are many. What are they? **(5)**

Q.3 ATTEMPT ANY THREE QUESTIONS: (15 MARKS)

(A) Write a short note on Active and Passive Metadata Repositories in a Data Warehouse. **(5)**
(B) Explain the term Taxonomy related to unstructured Data Warehouse. What are its types? **(5)**
(C) Explain the role of Total Information Quality Management stream with respect to seven stream approach to Data Warehouse. **(5)**
(D) What is meant by Heuristic Analysis of data in a Data Warehouse? **(5)**
(E) Differentiate Data Mart and Exploration Facility with respect to a Data Warehouse. **(5)**
(F) Explain Data Profiling and Mapping Stream it seven stream approach to Data Warehouse. **(5)**

Q.4 ATTEMPT ANY THREE QUESTIONS: (15 MARKS)

(A) Explain corporate data model with respect to a data warehouse. **(5)**
(B) Write a short note on Transformation of Data made as data passes from the Application / Interactive Sector to the Integrated Sector? **(5)**
(C) Discuss the relationship between Data Models and Unstructured Data. **(5)**
(D) What is the role of data quality monitor in a data warehouse? **(5)**
(E) Why is the Data Warehouse Monitor a standard recommendation for DW 2.0 environment? **(5)**
(F) Write a short note on Document Data. **(5)**

Q.5 ATTEMPT ANY THREE QUESTIONS: (15 MARKS)

(A) Differentiate between Discrete Time-Variant Data and Continuous Time Span Time-Variant Data. **(5)**
(B) How does data flow into the Integrated Sector of a Data Warehouse? **(5)**
(C) Data flow into near Line Sector of a data warehouse is considered optional. Why? **(5)**
(D) How is the term "Exception-Based Flow of Data" associated with flow of data in a Data Warehouse? **(5)**
(E) Discuss the points to be taken care of when source-to-target mapping of each unit of data has to be done to form a Data Warehouse. **(5)**
(F) Data throughput is a major concern with ETL processing. Justify. **(5)**

[TURN OVER]

Q.6 ATTEMPT ANY THREE QUESTIONS: (15 MARKS)

(A) How Indexing Technique helps to improve performance of a Data Warehouse? **(5)**

(B) How is migration of unstructured data different from navigation of structured data in a Data Warehouse? **(5)**

(C) Discuss the importance of good online response time to increase the performance of a Data Warehouse. **(5)**

(D) What are the different functions of the granularity manager in a Data Warehouse? **(5)**

(E) Capacity planning puts the organization in a proactive stance when it comes to performance. Justify. **(5)**

(F) Parallelization of processing is a really good way to enhance performance. Justify. **(5)**

Q.7 ATTEMPT ANY THREE QUESTIONS: (15 MARKS)

(A) Write a short note on an Optimizing Storage. **(5)**

(B) Explain with an example Data Clustering. **(5)**

(C) What are the techniques to improve the performance of data warehouse? **(5)**

(D) What is the advantage of using pilot systems? **(5)**

(E) Explain the procedure of loading dimension tables before the fact tables. **(5)**

(F) Write a short note on user acceptance procedure. **(5)**
